DOCUMENT RESUME

ED 459 834 IR 058 370

AUTHOR Nanard, Marc; Nanard, Jocelyne
TITLE Cumulating and Sharing End Users' Knowledge To Improve Video Indexing in a Video Digital Library.
PUB DATE 2001-06-00
NOTE 10p.; In: Proceedings of the ACM/IEEE-CS Joint Conference on Digital Libraries (1st, Roanoke, Virginia, June 24-28, 2001). For entire proceedings, see IR 058 348. Figures may not reproduce well.
AVAILABLE FROM Association for Computing Machinery, 1515 Broadway, New York NY 10036. Tel: 800-342-6626 (Toll Free); Tel: 212-626-0500; e-mail: acmhelp@acm.org. For full text: http://www1.acm.org/pubs/contents/proceedings/dl/379437/.
PUB TYPE Reports - Descriptive (141) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Access to Information; Archives; *Electronic Libraries; Indexes; *Indexing; Information Networks; Shared Resources and Services; *Users (Information)
IDENTIFIERS *Video Technology

ABSTRACT

This paper focuses on a user driven approach to improve video indexing. It consists in cumulating the large amount of small, individual efforts done by the users who access information, and to provide a community management mechanism to let users share the elicited knowledge. This technique is currently being developed in the "OPALES" environment and tuned up at the "Institut National de l'Audiovisuel" (INA), a National Video Library in Paris, to increase the value of its patrimonial video archive collections. It relies on a portal providing private workspaces to end users, so that a large part of their work can be shared between them. The effort for interpreting documents is directly done by the expert users who work for their own job on the archives. OPALES provides an original notion of "point of view" to enable the elicitation and the sharing of knowledge between communities of users, without leading to messy structures. The overall result consists in linking exportable private metadata to archive documents and managing the sharing of the elicited knowledge between users' communities. (Contains 22 references.) (Author/AEF)

Cumulating and Sharing End Users' Knowledge to Improve Video Indexing in a Video Digital Library

By: Marc Nanard & Jocelyne Nanard

Cumulating and Sharing End Users Knowledge to Improve Video Indexing in a Video Digital Library

Marc Nanard
LIRMM
161 rue Ada, 34392 Montpellier France
Phone: (33) 467 41 85 17, Fax: (33) 467 41 85 00
mnanard@lirmm.fr

Jocelyne Nanard
LIRMM
161 rue Ada, 34392 Montpellier France
Phone: (33) 467 41 85 17, Fax: (33) 467 41 85 00
jnanard@lirmm.fr

ABSTRACT

In this paper, we focus on a user driven approach to improve video indexing. It consists in cumulating the large amount of small, individual efforts done by the users who access information, and to provide a community management mechanism to let users share the elicited knowledge. This technique is currently being developed in the "OPALES" environment and tuned up at the "Institut National de l'Audiovisuel" (INA), a National Video Library in Paris, to increase the value of its patrimonial video archive collections. It relies on a portal providing private workspaces to end users, so that a large part of their work can be shared between them. The effort for interpreting documents is directly done by the expert users who work for their own job on the archives. OPALES provides an original notion of "point of view" to enable the elicitation and the sharing of knowledge between communities of users, without leading to messy structures. The overall result consists in linking exportable private metadata to archive documents and managing the sharing of the elicited knowledge between users communities.

Categories and Subject Descriptors

H.3.5[INFORMATION STORAGE AND RETRIEVAL]: Online Information Services - *Data bank sharing*

General Terms

Design

Keywords

Video annotation. Video indexing. Private workspaces. Users communities. Knowledge sharing.

1. INTRODUCTION

It is now well admitted that retrieval of relevant images or video segments among large collections requires taking advantage of semantically rich metadata associated to small information chunks. A lot of efficient techniques for automatically elaborating metadata from text documents are now well mastered. References on that topic can be found, for instance, in conferences on information retrieval [19].

Conference JCDL '00, May, 2000, Virginia.

At the opposite, automatically elaborating semantically relevant metadata from images and, moreover, from video is a far harder task [1] which currently is a challenge for further development of the information technologies and multimedia digital libraries. The cause is obvious: contrary to texts which, as a natural language representation, have the power and all of the features of a formal knowledge representation scheme, images only rely on an iconic representation scheme [18]. They rely only on suggestive, emotional communication modes. They do not usually embed any syntactic or semantic structures likely to be elicited by a machine for elaborating semantically rich metadata. As a consequence, and unfortunately, human interpretation of video still is the only one technique which enables precise semantic indexing at scene level.

Automatic image indexing techniques have huge difficulties in accessing the semantics of an image. The simplest image indexing techniques do not care at all for image semantics. They are based on signal processing. They focus only on physical and graphical properties of the image [3] such as the color histogram, the textures, image similitude, and so on, without any interpretation. A more elaborated approach takes advantage of image recognition. Such techniques currently remain limited to simple cases such as very typical faces recognition [5], [13], situation recognition (sitting/standing), familiar object recognition (cars, planes, tables). Nevertheless, very little semantics can be elicited from image analysis. A far more efficient approach consists in taking advantage of multimodality between image and sound tracks in movies or in TV news broadcast for cross fertilizing the document analysis. In the Informedia project [12], [17], the recognition of a subset of relevant words such as politicians or country names in the sound track of news may let attach, for instance, to a landscape image a metadata telling that the image concerns "Afghanistan", since this word has been recognized in the voice commentary. This technique also helps contextually solving ambiguities in image recognition. For instance, let us suppose the system recognizes the presence of a face but cannot identify it further. Famous names recognition in an associated commentary on the sound track may help the system improve recognition, solve ambiguities and let it suppose it is, for instance, Marilyn Monroe's face. This quite efficient technique for automatically indexing news is already available on the market place. Nevertheless, none of these automatic techniques can fully succeed in automatically indexing a large variety of archive documents. Either there is no or too few associated multimodal data, or the commentary is only very loosely related to the image, like this is unfortunately the case in many news report. Therefore only the mixing of several approaches can lead to a better indexing of images and of videos. In most cases, correct indexing of images and of video requires human interpretation of the situations.

In this paper, we focus on a typically different approach to improve video indexing. The approach does not intend at all to be a substitute to other. Rather it is a complementary strategy for drastically improving the overall efficiency of the end user's work in the trend of social navigation [7], [8]. It consists in cumulating the large amount of small, individual efforts done by the users who access information, and in providing a community management mechanism to let users share the elicited knowledge. This technique is currently being developed in the OPALES environment and tuned up at the Institut National de l'Audiovisuel in Paris (INA) to increase the value of its video archive collections. It relies on a portal providing private workspaces to end users, so that a large part of their work can be shared between them. The effort for interpreting documents is directly done by the expert users who work for their own job on the archives. OPALES provides with an original notion of "point of view" to enable the elicitation and the sharing of knowledge between communities of users, without leading to messy structures. The overall result consists in linking exportable private metadata to archive documents and managing the sharing of the elicited knowledge between user communities.

The paper first describes the context of the study and its design rationale. Then it focuses on a specific point of the project: the management of user elicited knowledge. The notion of "point of view" enables to reduce the problem complexity. It helps manage smaller knowledge clusters specific to user communities.

2. CONTEXT OF THE WORK

In any domain of industry, companies usually keep track of their own production, most often for technical or commercial reasons, but sometimes also as archives considered as a memory of patrimony. We name these kinds of archives "patrimonial archives". For instance, car producers build large museums to exhibit tracks of their creative activity. In any cases, these archives represent a very small part of their production. Contrary to goods manufacturers, information producers deal with such a huge amount of data that keeping all of it for a long time is a hard and costly choice. Whereas policies for archiving printed documents for the long term are now ruled at national level in many countries, video production is not yet concerned with such rules. Storage is often handled directly by producers, and thus storage strategies may be subject to opportunistic variations. As a consequence, a large part of TV production is discarded once it has been broadcast. In many cases, just the best part, or the reusable part is preserved. Even, in TV or radio companies where systematic archiving is often the rule, heavy storage cost, lack of room for storage, inconsistencies in the storage strategy or changes in the management sometimes lead to later discard archives which had been preserved for years. For instance, such a situation had already occurred, leading a few years ago a famous broadcasting company to discard a large part of its records of daily news of the fifties.

2.1 INA, multimedia archive provider

The Institut National de l'Audiovisuel (INA), created in Paris in the early seventies, is in charge of keeping records of national French TV broadcasts. A law voted in June 1992 defines the "*dépot legal* (official and mandatory storage)" which requires copies of any national radio or TV production to be deposited at INA as patrimonial archives. Storage does not concern simply the items themselves (e.g.: TV series as such) but also the context in which they have been broadcast. This enables rich sociologic studies, for instance studies of correlation between the focus of advertisements and the contents of the film they break. Similarly, the context associated to the audio and video contents provides historians with a far more precise record of our way of life than separate items would do. Furthermore, INA has inherited from the archives of the previous national broadcasting company "ORTF". Currently, INA deals with more than one and a half billion of hours of TV and radio and more than one billion of still pictures stored on more than fifteen miles of shelves. INA already has started to convert its data to digital format. 200 000 hours of TV and 300 000 hours of radio are now available, thus making it the repository of one of the largest collection of audio-video archives, like those of BBC and RAI.

INA's main function is to be an information provider for TV producers, and for any other professionals. INA is famous in France for its authentic and watermarked archive sources included in TV news. It also serves as a patrimonial archive library for researchers such as historians, sociologists, economists, politicians, and so on, who study historical facts. Since INA is just the archivist but is not the copyright owner of all of deposited archives, it often operates just as a partner between buyers and information owners.

Efficiently accessing such a huge amount of archives is an increasingly important challenge for INA. Like in any library, the video archives have been indexed once for all when they were stored. This initial indexing is obviously sufficient for most of professional use: everyday TV producers access the INA video-library to search and buy archive sequences. Of course, it is not possible, nor suitable to make changes in this primary indexing scheme to improve it.

One way to offer better service to users is to build a new separate indexing, based on more efficient and more precise techniques such as NCG [4], enabling video indexing at different levels of granularity and stratification of indexing [20]. Unfortunately, the cost for re-indexing the entire set of archive documents is far beyond the possibilities of the organization. So, the planned solution is to let it be done by the users themselves and to incite them to cumulate their individual efforts to improve the overall service.

2.2 The OPALES project

2.2.1 Overview

OPALES is an ongoing R&D project, initiated by the French ministry of Economy in 2000, scheduled to be operational in the fall 2001. It aims at developing a new service empowered by digital video and hypermedia technology, and intended to incrementally increase the value of the multimedia archives accessed through it. It consists of a distributed environment able to support the activity of virtual communities of experts working on the INA patrimonial video archives. OPALES is a private portal. It enables its users to directly work on archive documents in private workspaces, to share elicited knowledge about studied documents, and to collaborate anonymously as well as within explicit groups. The basic assumption is that the results of the work of expert groups can be made available to others, thus boosting their own work. The return of business generated by knowledge exchange between experts is also business for the archive provider itself.

2.2.2 *Target users*

Access to the OPALES portal is currently restricted to a group of researchers who participate to the R&D project. Beside INA, several institutions participate to its elaboration and evaluation: the "Cité des Sciences et de l'Industrie" in Paris, the MSH "Maison des Sciences de l'Homme", the CNDP "National Center for Distance Learning", and the BPS "Program and Service Bank" of the 5th TV Channel. They provide expert users as well as video data. The targeted users are typically knowledge workers. For the first steps of the project, researchers in human sciences and teachers have been chosen as representatives of future users of the system. They access documents and study them with the purpose of elaborating new knowledge, either for their own usage or for transmitting it to others.

2.2.3 *Corpus*

In order to make experiments easier and cheaper, the corpus currently used to bootstrap the project only contains copyright free documents. Handling copyright issues is of course one of the usual INA business. But this point is beyond the scope of the first stage of the project.

2.2.4 *Task*

The task supported by OPALES is called "active reading". Researchers usually practice active reading in libraries. They act as readers and writers at the same time. They annotate, extract, search, etc. Such a task consists of alternated reading and writing steps deeply intermingled, thus producing a gloss bound to the document. Although the term "active reading" had been coined for working on printed documents, this task also concerns video documents. Actively reading a video is fundamentally different from simply "looking at" it. It supposes the will to understand the document in its depth, to connect facts with others, compare sequences, and so on. To do so, the reader needs to create private notes, to link them directly onto segments of the read video, exactly like a researcher annotates a private copy of a paper. Active readers also frequently wish to know what other readers think about the studied documents. Of course, the reader is usually an author who writes her own documents, inserts archive items into them, and annotates them in the same manner. For instance, a history teacher at a university enjoys preparing her own video from highly relevant archive segments selected to illustrate her discourse.

All of these considerations make the INA portal quite different in its purpose from portals of most of Internet access providers.

3. DESIGN RATIONALE OF OPALES

The OPALES project relies on the following assumptions:

- Sharing one's knowledge with other people improves one's work efficiency [22].
- One uses a tool only when the return is greater than the effort to use the tool.
- To be efficient on a machine, a user needs interacting seamlessly with the objects (s)he studies as well as those (s)he produces.

To do so, OPALES provides each of its registered users with a private workspace. The purpose of the workspace is threefold:

- Enable the user to work on archive documents and on other documents as freely as if they were private copies, and to use them as raw material for their own use.
- Keep track not simply of the "production", but also of the work, e.g. the interpretation of facts observed on the videos. We call it "elaborated knowledge".
- Manage the sharing of elaborated knowledge with other users. This last point implies the use of efficient but flexible open collaboration techniques in order to facilitate structure emergence from the end users efforts [7].

The overall result is also threefold:

- The user produces for her own use new documents and new knowledge from the archives. This is supposed to be the basic reason why (s)he works on the system. No one sustains a long effort when there is no personal return.
- The effort done by a user at work is capitalized by sharing it with others. This results in a direct return from the OPALES system which incrementally improves the available knowledge about documents.
- Knowledge sharing between users can be done either for free or be accounted, in this case generating knowledge business. Some expert group may import knowledge about the archive documents from other expert groups to improve their own understanding of documents and provide other experts with this improved knowledge. Dealing with knowledge business is out of the current scope of OPALES whereas knowledge sharing accounting is already handled in the system.

These considerations match the initial goals:

- First, capitalizing and sharing user knowledge in the system boosts everyone's efficiency. This idea was strongly promoted by Douglas Engelbart. One may consider OPALES as an implementation of a NIC (Network Improved Collectivities) [9].
- Second, the result of users work directly benefits to the owner of the portal: the elaborated and capitalized knowledge constitutes an added value to the documents, which makes them more attractive and more valuable for access by new users through the portal.
- Third, users access documents on the OPALES portal for working and preparing their own documents. The workspace offers seamless interaction with any kind of document: from archives documents to users' own documents and even to documents built as shared knowledge.

4. THE POINT OF VIEW NOTION

4.1 Design rationale of the point of view notion

4.1.1 *A shared ontology*

Sharing knowledge implies that the users agree on the meaning of some vocabulary. This is done by representing knowledge in the system according to a shared ontology [10], [11]. This ontology is used internally in OPALES for indexing documents and computing on indexing.

Nevertheless, two major problems must be solved for cumulating user efforts:

- Providing users with an extensible representation mechanism for freely representing their own knowledge.
- Inducing a strong structure of the resulting knowledge in a non intrusive way.

4.1.2 Extensible Ontology

The first problem implies that the ontology cannot be static. Although OPALES is a restricted access system open to people who share the same need to understand and interpret archive contents, there is no restriction on the topics on which experts focus. Moreover, the diversity of expertise domains is precisely the interest of the system, because no library could afford such a large panel of experts to index the documents.

When annotating video sequences, experts in a given domain need to be allowed to handle concepts specific to their domain, which are mostly specialization of existing ones. As a consequence, they must be allowed to enhance the shared ontology accordingly, under some control.

4.1.3 Non intrusive interaction scheme

The second problem implies finding a good balance between constraints and freedom. This is one of the originalities of OPALES. If the structure is too strongly constrained by the system, in an intrusive manner, the user in hampered. Her activity reduces and the overall efficiency collapses. Conversely, if the structure is too weak, the knowledge elaborated by some users may become soon incompatible with the knowledge elicited by others, leading to messy and unusable results. As a consequence, regulation mechanisms based on community management are needed to avoid an anarchic evolution of the ontology. This mechanism is provided in OPALES, owing to the choice of an internal knowledge representation scheme directly computable. It enables the system to control for example the evolution of the ontology and to make users who edit the ontology aware of the existence of concepts similar to those they want to add.

4.1.4 Points of view as knowledge clusters

To deal with these problems, OPALES introduces the original notion of "point of view" which enables to virtually organize the users work into dynamically adaptable virtual communities in order to manage clusters of locally consistent knowledge. Dealing with inconsistency is a complex and delicate problem, even for humans. It becomes harder and harder as and when the scope of the knowledge widens and the amount of metadata increases, which is the case in OPALES. In order to keep the inconsistency in reasonable and manageable limits, we have made the choice to break it down, by dynamically identifying smaller scopes of knowledge in which sets of users can locally manage by themselves the consistency of their sub-domain. The result is that knowledge is self-organizing in locally consistent small clusters which directly reflect the structure of user expert groups. For instance, if some users have expertise in "fashion and dressing in the sixties" and need to introduce new concepts in the ontology, it is easier to them to locally manage the suitable extension. Thus, evolution of the ontology remains local and does not conflict with extensions needed by other experts, for instance those of "horses races". In order to insulate the clusters and organize their overall structure, a technique similar to XML namespace is used: we call it a "point of view". The extensions of the ontology and of the elicited knowledge are explicitly attached to the domain for which they have been added: they belong to a "point of view".

OPALES provides means to create at will clusters called "authoring points of view" and to elicit knowledge into them. It symmetrically provides means to take advantage of knowledge elicited according to different points of view, so that a reader may mix the knowledge elaborated by several communities.

4.2 Virtual communities

Most of OPALES users are experts, for instance in history, sociology and so on. Their expertise makes them, implicitly or explicitly, belong to "virtual communities". A community is said virtual when its members do not need to know each other. A virtual community exists as soon as some people have identified and named their concern, thus making explicit to others some knowledge, some interest, some hobby, and wish to share it, anonymously or not, with others [15], [6]. Virtual communities emerge on the web everyday. We call such communities virtual to stress the fact that belonging to a community does not require to be introduced, to pay for it, nor to adhere to some predefined ideas. A virtual community exists when a topic is made explicit by naming it and precisely identifying it, and when some people feel concerned by it. In OPALES, a virtual community is implicitly created when an author defines a new point of view and makes it public. At that moment, other users can feel concerned with writings related to this point of view as readers or as authors.

4.3 The notion of "point of view" in OPALES

4.3.1 Definition

The term "point of view" seems quite familiar but is used in OPALES with a very precise and restrictive meaning. We define it as a statement of the author about her authoring activity which sets the document in the concerns of a virtual community. Contrary to some familiar meaning, the "authoring point of view" of a document is not the semantics of the document itself. For instance, two experts may annotate a video on "Cashmere War" with completely contradictory interpretations, whereas they share a same vocabulary to express it, and have the same concern. In OPALES, their annotations belong to the same point of view: "India and Pakistan matters experts" regardless to the actual content of the annotation. Conversely, the same video may be annotated with the point of view of a "video reporter school teacher" who would comment the narrative structure, the framing of shots, the choice of images and so on. "India and Pakistan matters" and "video reporter school teacher" are quite distinct points of view. They can be used to annotate the same document. A "Economical international relationships expert" would annotate the same document in a quite distinct manner.

The notion of point of view in OPALES enables writers to explicitly tell to which virtual community their writings are dedicated. It induces clustering of knowledge and enables to use the specific community vocabulary which is appended to the shared ontology as depending on the point of view. It implicitly defines in this way local namespaces which drastically reduce ambiguities.

4.3.2 Using and managing points of view

The kernel of OPALES internal architecture handles private and public documents, points of view, annotations and indexing in a unified, reflexive, and consistent manner. Consequently, we use the term "piece of information" rather than the term "document" which could be understood with some restrictive meaning. To any piece of information is attached a resource descriptor which includes an "authoring point of view" stamp, an owner stamp, a type, and a status tag, and so on. For portability reasons, resources are externally described as RDF descriptors [21], [14]. A "workspaces database" keeps track of all the resources and of their interdependencies. Points of view are implemented like stamps attached to any piece of information. They characterize in

which context information makes sense. For reflexivity reason, points of views are also considered as "pieces of information": a unique document of type "point of view" (which is primitive in the system) is associated to each point of view, as its informal description. This document is indexed by a precise indexing pattern, which enables the system to retrieve points of views. Thereby, there is strictly no difference between indexing points of view and other documents. The same mechanism applies for retrieving them.

The role of this mandatory indexing pattern associated to each point of view is to formally characterize it with respect to the shared part of the ontology from which the point of view is visible. It enables any author both to retrieve existing points of view defined by other authors and to declare new ones so that other authors can be aware of their existence. For many reasons, which are out of the scope of this paper, the OPALES internal knowledge representation formalism is NCG, the "nested conceptual graphs model" [4]. NCG enables a more precise indexing than keywords. For instance, NCG makes it very simple to distinguish between "transportation of sailing boats", "transportation by sailing boat", and "transportation of sails of boats". Another important result about NCG is a fuzzy matching algorithm [16] used for comparing NCG representations; it takes advantage of specialization, generalization and composition relationships in the ontology. It enables to compute distances between NCGs and thus to determine which are the closest points of view to a given one. For instance, an expert analyzing a movie of the 2nd World War can annotate it from a "medical expert" point of view or from one of its specialization as "nutrition expert" or as "psychiatry expert". As a consequence, the search engine would retrieve psychiatry annotations as specialization of medical expert annotations. Points of view and vicinity of points of view are the base for retrieving annotated documents and annotations, which are meaningful for a virtual community. This is the internal basement for the points of view and virtual community management in OPALES.

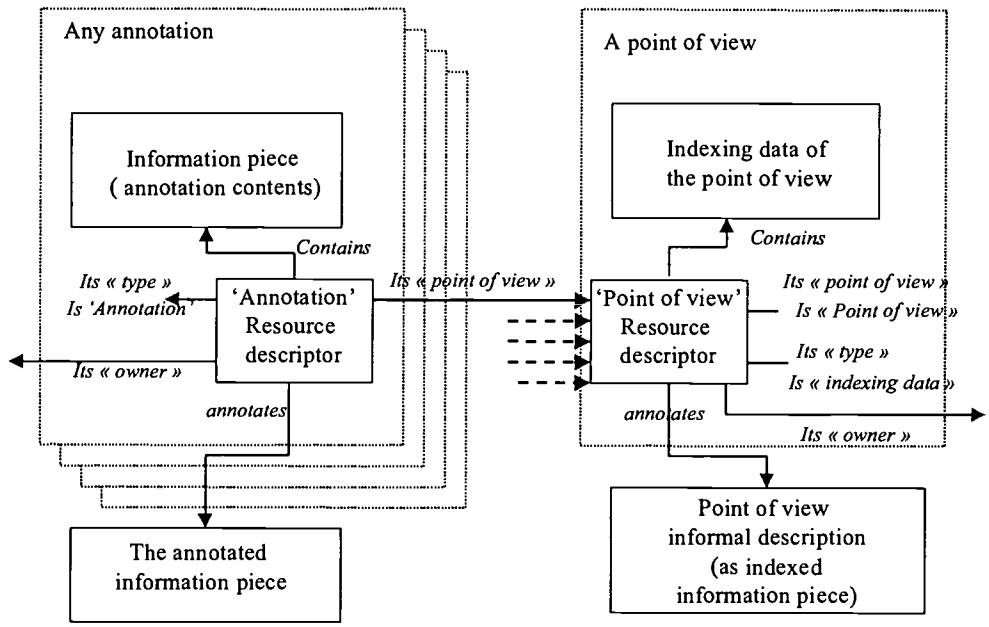

Figure 1:

Reflexivity in OPALES internal structure: annotations, indexing, points of view... are handled in a unified manner.

4.4 How authors interact with points of view

4.4.1 Selecting or defining a point of view

One of the requirements of OPALES design is a very low overhead for users. The point of view management sub-system is designed so that it provides users with more return than it requires efforts to put it in action. Any created piece of information (annotation, document, indexing) automatically becomes a resource stamped with the point of view associated to the window in which it was edited, and typed by the editor's type.

When a user logs in OPALES, her private workspace displays the last state in which the user logged out. Thereby, the list of her favorite authoring points of view, as created in previous sessions, is already available. A "current" point of view is kept marked in the list. It is assigned to any new window for stamping any editing actions taking place in it. A pop up menu enables to easily change the "current" point of view of a window whenever needed.

As for any other document, retrieval of a point of view not in the favorite list is achieved by means of a query. OPALES interface helps elaborating the query according to the ontology, by contextually selecting the vocabulary. Points of view close to the favorite ones can also be directly accessed in a browser interface. If the user considers that no existing point of view matches her current authoring situation, she creates a new one, most often by specialization of an existing one. Let us remark that, if no relevant point of view can be found, the query itself is very close to the formal indexing of the new point of view, thus making the burden to create new points of view quite limited.

All this just requires the author is conscious of the context in which she works. This assumption is fully compatible with OPALES users groups.

In most of cases, annotating existing documents or creating new ones does not require the author explicitly deals with points of view, since the current point of view is automatically assigned by default when an information chunk is created.

4.4.2 Exporting points of view

Any information piece (or document) in OPALES has a status tag which indicates whether the chunk is public or private. A private document can be accessed only by its author, whilst a public document can be read by anyone but edited only by its author. For consistency internal reasons and use of reflexivity in the implementation architecture, points of view are handled as documents. For sure they are so, because they have a content (their informal description), they are indexed exactly like any other document, they have an author who created the point of view, and a point of view ("point of view creator" which is primitive in the system). As a consequence, like any document, a point of view can be either private or public. Making a document or a point of view public is called "exporting" it. This makes it potentially visible to other users. This enables users to privately handle their annotations in their private workspace and later export them as well as the associated points of view.

4.4.3 Owners of documents

Any piece of information resource in OPALES has an owner and a point of view. No one except its owner may edit a piece of information. For consistency reasons, this applies to archive documents as well as to annotations and private documents. The term owner must be understood not as the copyright ownership but as the person or the institution who is responsible of the storage of the information in the system. An archive (video, image, sound record, text...) is under the responsibility of an institution (INA, MSH,...) who added it to the portal ; the institution is its OPALES "Owner". The point of view of an archive document is "archive" which is primitive in the system. This is quite consistent with the notion of point of view: for instance, an indexing with the "archive" point of view precisely is the genuine "INA" indexing associated to the document. Like any other document an archive can be public or private. In this last case, it is not visible for the end users, but may be handled by its owner. This feature is useful for instance during the first indexing stages of documents done before exporting them.

4.5 Annotating videos with OPALES

4.5.1 Stratified annotations

OPALES allows stratified [20] indexing and annotation of video. Freely stratified annotations are independent annotations whose anchoring in a document may overlap at will. Although automatic scene recognition tools easily provide a primary segmentation of video, it is now well known that this kind of segmentation is insufficient for precise indexing. For instance, in news, topics are announced and start with the speaker face on the screen. Automatic scene separation suggests starting a new segment when the image changes from the speaker to another image, whereas such an event may occur in the middle of a sentence. Breaking it or shortening it may deeply alter its semantics. This kind of segmentation is visual but, not at all, semantic, like those which are the concerns of OPALES. Because users index and annotate documents themselves, they are allowed to freely define segments and annotate them. For instance a specialist of body language may study hand motion of politicians during speeches. The segments she needs in order to put her expertise in action are quite different from those needed by a specialist of rhetoric. Stratified indexing is suitable so that annotations can freely overlap.

4.5.2 Annotation versus indexing

An annotation is an informal metadata, i.e. any information piece linked to a document. In OPALES there is no constraints on its content. An annotation can be simply the name of a person on an image of a group of guys and a link with a geometrical anchor to locate the person on the image. It may also be a long and argued discussion about some events of the currently selected segment. It can be a typed link towards another document.

At the other extreme, indexing is a formal data anchored into a document, and internally represented as a NCG. Formally indexing a document consists in providing typed annotations (type is "indexing", which is primitive) containing computable metadata which enables the internal search engine to retrieve it. Since indexing is just a specialization of annotations, as many private indexing, with specific points of view can complement the archive indexing of a document and thus describe richer semantics on specific segments as well as on the whole document.

Indexing a video segment or any part of a document is achieved by making a selection in the information piece and opening an annotation window of type "indexing". A specific NCG based indexing tool opens in the annotation windows. Indexing patterns can be defined by communities of users and attached to points of view in order to help indexing and ensure consistency of indexing rules within a point of view. Regulation mechanisms are provided by the user community management sub-system. Some virtual groups may become explicit, work closer together and elect moderators. This is a problem of user management, which is out of the scope of the paper.

5. EXPLOITING THE NOTION OF POINT OF VIEW

5.1 Reading versus authoring points of view

The notion of point of view would have no interest if it were not the key feature for readers working on documents. It is used to improve the information retrieval mechanism and provide finer access to the annotation base. We distinguish the notions of "authoring point of view" and of "reading point of view".

On the one hand, an authoring point of view characterizes the virtual community dedicated by an author to an annotation when he creates it. An annotation or an indexing is characterized by only one authoring point of view. On the other hand, a reading point of view characterizes which sources of annotations a reader wants to see linked as complements to a displayed document, and which complementary indexing information the OPALES search engine will use to retrieve more relevant documents. A reader can use different reading points of view to observe annotations and indexing of video segments.

Therefore, authoring points of view and reading points of view are distinct notions handled separately by the system. Let us suppose a reader wishes to integrate sociologic and economic sources as complementary information in her studies in order to

get a deeper understanding of the studied videos. For retrieving more relevant videos, she also mixes in the queries concepts defined in extension on the ontology part associated to these points of view. The union of "economy" and "sociology" corresponds to her "reading point of view". Her authoring point of view simply is "childhood expert" which is her specialty. She considers her neither as a sociology expert nor as an economy expert and would not write annotations or indexing as such. She imports these points of view in her workspace just to constitute a "reading point of view". She may export her annotations written with the "childhood expert" point of view, inducing in this way a kind of knowledge commerce between users.

5.2 Defining a reading point of view

In a user's workspace, any editor or browser window has an associated "reading point of view" which acts as a filter to enhance its contents. The favorite reading points of view of a user are kept in a list in order to enable her to quickly set the point of view associated to her windows. Defining a new reading point of view is usually achieved by specifying an ordered set of authoring points of view. The reader just drags and drops some authoring points of view to define this new reading point of view. She can explicitly name it for further reuse. She can also explicitly define it in the same manner as a new authoring point of view, for instance by taking advantage of generalization mechanisms.

A list of annotations selected according to the reading point of view associated to a window is dynamically associated to the currently displayed document. The listed annotations are those which have been authored in one of the points of view referenced in the reading point of view, and which were linked as annotations anchored to the current selection in the displayed document. For instance, let us suppose the reader has selected some segment of an archive video as an answer to a search query, and looks at it. Since she observes it from a given reading point of view, all the available annotations for this point of view which are linked to any segment of this video that includes the current time code are listed. When seeking the video, the annotation list is dynamically updated according to the current position. Moving the cursor over the list displays a short preview of the selected annotation, thus avoiding unnecessarily link firing. When an annotation is geometrically anchored into the video, moving the mouse over its reference in the annotation list shows its anchorage directly on the video, under the condition the video is in the paused mode. This feature is extremely pleasant, for instance for scanning names of participants on a picture of a group.

5.3 OPALES system architecture

OPALES system architecture, as shown on figure 2, relies on the cooperation of three servers. The main server delivers archive video data and icons of selected shots. The workspace server stores all private and shared information pieces which are not archives, and uses a database for managing descriptors. It delivers enhanced information according to the selected reading point of view. Most of interactions are locally handled by a plug-in on the client browser. The knowledge server is based on a NCG engine developed at LIRMM [16]. It stores the ontology and all the indexing data.

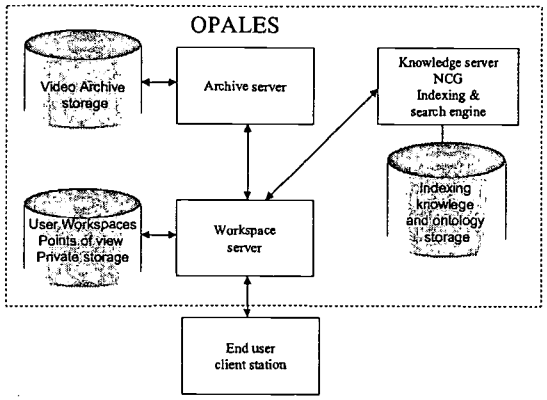

Figure 2: OPALES system architecture.

6. DISCUSSION

The structure of users' work with OPALES emerges as the consequence of using a very simple set of rules associated with the private workspaces:

- Each user feels like working privately on her own copies of documents.
- If a reader selects the "archive" point of view, she only sees genuine information.
- If a reader imports some points of view, the displayed documents are enhanced with annotations accordingly.
- Searching for points of view is done in the same manner as searching for documents.
- Only the owner of an information may alter it. Imported information is inalterable.
- All information pieces created by a user keep track of the point of view in which they were created.
- A user may export and import points of views.

As a consequence,

- Any information made public is always, *de facto*, organized into a structure based on the point of view description in the ontology. When it is exported, it is cumulated in the system in an organized and non intrusive manner for the users, which induces very little overhead.
- The cumulated effort is made available to the collectivity of users in such a way that a user may focus only on her sub-domains. The reading point of view acts as a dynamically adjustable filter, which spares the burden to express complex queries. Furthermore, the point of view notion is far richer to express semantics than keywords are, since it precisely expresses the author's intention, whether or not relevant keywords are present in the annotation.

7. CONCLUSION

Patrimonial video archives contain considerable amounts of highly valuable information about our society. Contrary to books, which can be automatically analyzed once digitized for enhancing their indexing, digital video still requires human expertise to be relevantly indexed. The OPALES project offers a solution to enhancing the elicited knowledge about a part of the INA archive library.

Relying on users' work is a challenge. The web has assessed the outstanding power of users collaborating together. The Semantic Web Project [2] trusts this assumption as well. OPALES design aims at providing users with both simple and efficient mechanisms to share their knowledge. Ease of use seems to us a strict prerequisite to bootstrap knowledge sharing between users, and to cumulate it in the library. The concept of "point of view" and its implementation in OPALES are a key for reducing the complexity of huge amounts of knowledge independently elicited by groups of users. Although OPALES has been designed for enhancing video archives, the described techniques are directly transposable to other types of digital libraries.

Experiments for observing users' behavior and adjusting mechanisms are on the way.

REFERENCES

[1] Aigrain, P., Petkovic, D., & Zhang, H.J. Content-based representation and retrieval of visual data: a state of the art review, Multimedia Tools and Application, Special issue on representation and retrieval of visual media, 1996.

[2] Berners-Lee, T. Semantic Web Road Map, http://www.w3.org/DesignIssues/Semantic.html, 1998.

[3] Chang, S.F. *et al.*, VideoQ: An automated content-based video search system using visual cues. In Proc. ACM Multimedia'97 (1997), pp. 313-324.

[4] Chein, M., Mugnier, M.L., & Simonet., G. Nested Graphs: A Graph-based Knowledge Representation Model with FOL Semantics, in Proc. 6th International Conference on Principles of Knowledge Representation and Reasoning (KR'98), (1998), pp. 524-534, Morgan Kaufmann Publishers.

[5] Crowley, J.L. & Berard, F. Multi-Modal Tracking of Faces for Video Communications, IEEE Conference on Computer Vision and Pattern Recognition, CVPR '97, Puerto Rico, (1997).

[6] Davenport, G. & Pan, P. I-Views: a Community-oriented System for Sharing Streaming Video on the Internet, in Proc. WWW9 Conference (1999).

[7] Dieberger, A. Supporting Social Navigation on the World-Wide Web. International Journal of Human Computer Studies: Special Issue on Novel Applications of the WWW (in press).

[8] Dieberger, A., Dourish, P., Höök, K., & Wexelblat, A. Social Navigation: Techniques for Building More Usable Systems, Interactions, Vol. VII.6, 2000.

[9] Engelbart, D., Networked Improved Communities, Keynote at ACM Conf. Hypertext'98, (1998). ACM SIGWEB video archive. See also http://www.bootstrap.org

[10] Garino, N. Formal ontology, conceptual analysis and knowledge representation. Int. Journal of Human-Computer Studies, 43 (5/6), pp. 625-640, 1995.

[11] Gruber, T.R. Toward principles for the design of ontologies used for knowledge sharing. In *Formal Ontology in Conceptual Analysis and Knowledge Representation*, Nicola Guarino and Roberto Poli, editors, Kluwer Academic, in preparation. Original paper presented at the *International Workshop on Formal Ontology*, March 1993. Available as Stanford Knowledge Systems Laboratory Report KSL-93-04. On line: http://ksl-web.stanford.edu/knowledge-sharing/papers/onto-design.rtf

[12] Hauptmann, A. and Smith, M. Text, Speech, and Vision for Video Segmentation: The Informedia Project, AAAI Fall 1995 Symposium on Computational Models for Integrating Language and Vision, 1995. See also: http://www.informedia.cs.cmu.edu/

[13] Houghton, R. Named Faces: Putting Names to Faces. IEEE Intelligent Systems Magazine, Vol 14, No. 5, pp. 45-50, 1999.

[14] Kahan, J., Koivunen, M.R., Prud'Hommeaux, E., & Swick R.R. Annotea: An Open RDF Infrastructure for Shared Web Annotations, in *Proc. of the WWW10 Int. Conference*, Hong Kong, (2001).

[15] Martin Röscheisen, M. & Winograd, T. Beyond browsing: shared comments, soaps, trails, and on-line ccommunities, in Proc. WWW3 Int. Conference (1995).

[16] Mugnier, M.L. Knowledge Representation and Reasonings Based on Graph Homomorphism, in Proc. 9th International Conference on Conceptual Structures (ICCS), (2000).

[17] Olligschlaeger, A., Hauptmann, A. Multimodal Information Systems and GIS: The Informedia Digital Video Library, ESRI User Conference (1999).

[18] Peirce, C.S. Ecrits sur le signe, Editions du Seuil, Paris, 1978.

[19] Staab, S., Erdmann, M., Maedche, A., & Decker, S. An Extensible Approach for Modeling Ontologies in RDF(S), Workshop on Semantic Web associated to ECDL'2000.

[20] SIGIR conferences, ACM Press.

[21] Smith, A., & Davenport, G. The Stratification System: A Design Environment for Random Access Video. In ACM Workshop on Networking and Operating System Support for Digital Audio and Video, San Diego, California (1992).

[22] Stone, V.E. Social Interaction and Social Development in Virtual Environments. Presence 2, 2 (Spring 1993), pp. 153-161.

U.S. Department of Education
Office of Educatonal Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

REPRODUCTION RELEASE
(Specific Document)

NOTICE

REPRODUCTION BASIS

☒ This document is covered by a signed "Reproduction Release (Blanket) form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").